Environment. Water Quantity Management Branch.

Ontario

Ministry of the Environment

Hon. William G. Newman, *Minister*
Everett Biggs, *Deputy Minister*

Water Resources Paper 7

c. 1

Daily Streamflow Simulation for the Thames River Basin

INTERNATIONAL HYDROLOGICAL DECADE · DECENNIE HYDROLOGIQUE INTERNATIONALE

WATER RESOURCES
PAPER 7

DAILY STREAMFLOW SIMULATION FOR THE THAMES RIVER BASIN

By

S. Singer

MINISTRY OF THE ENVIRONMENT

WATER QUANTITY MANAGEMENT BRANCH

TORONTO ONTARIO

1974

PREFACE

As part of the Ministry of the Environment's hydrologic research studies, the River Basin Research Section is developing mathematical models to describe the complex interactions of the various components of the hydrologic cycle in research watersheds. During model development, it is useful to apply developed techniques to existing hydrologic questions in order to test the viability of these techniques for solving problems related to local hydrologic conditions.

The results of the application of the streamflow generator described in this study will form part of an overall comprehensive report on the Thames River basin management study, to be released at a later date. The development and application of this daily streamflow simulator is outlined separately here to indicate the use of such a technique possibly as an aid to others engaged in the simulation of streamflows.

K. Symons

K. E. Symons, Director,
Water Quantity Management Branch.

Toronto, January 1, 1974.

TABLE OF CONTENTS

APPENDIX I

LIST OF TABLES

APPENDIX I

LIST OF TABLES

APPENDIX II

LIST OF FIGURES

FOREWORD

In generating stochastic sequences of flow for the Thames River stations, using statistical properties of historic data, it is extremely important to realize that the sequences are generated by statistical methods that do not pretend to provide cause-effect models for actual flows. Fiering and Jackson (3) indicate that the data generated by using such methods "while neither actual historic records nor predictions of future flows, are close enough to possible (but not observed) historical records that may be used to determine, within defined statistical errors, the several quantities of interest," from the point of view of a water management study.

ACKNOWLEDGEMENT

The author wishes to acknowledge the co-operation of Environment Canada in supplying various computer programs for the study, in particular the assistance of D. R. Cuthbert, whose advice proved invaluable.

Thanks are due to B. J. Izatt of the Systems and EDP Section, Ministry of the Environment, for assistance in the computer application, and to F. C. Fleischer, Program Engineer with the River Basin Research Section of the same Ministry, for his guidance during the carrying out of the study and his critical review of this text.

DAILY STREAMFLOW SIMULATION FOR

THE THAMES RIVER BASIN

INTRODUCTION

In participating in the Thames River basin water management study, the Water Quantity Management Branch is presently undertaking several studies dealing with the collection, compilation and analysis of ground and surface water data. In order to provide essential input to the water quality simulation model under development by the Water Quality Branch, the Water Quantity Management Branch was requested to provide extended daily streamflow data sequences for several stations on the Thames River.

The objective of the study was to derive a suitable multiple-station daily streamflow generator for the Adelaide, Ealing and Byron stations in the vicinity of the City of London, using an acceptable mean monthly generator to provide the proper required input of synthetic monthly means to the daily flow generator.

THE THAMES RIVER BASIN

The Thames River, the headwaters of which are located in the highlands of the counties of Perth and Oxford, northeast of London, drains approximately 2,250 square miles. The river carries water for more than 190 miles to Lake St. Clair. The Thames River above London (Upper Thames) has numerous tributaries, all of which drain into two main branches: the North Branch (app. 660 sq. miles) and the South Branch (app. 530 sq. miles), the confluence being within the City of London. Below London, the Thames River (Lower Thames) flows through a comparatively narrow drainage area where all the tributaries appear relatively-small (9) (see Figure 1, Appendix II).

The mean annual temperature in the basin ranges from 48.7°F at Chatham, near Lake St. Clair, to 43.8°F at Mitchell in the Upper Thames. The mean annual precipitation increases generally in an upstream direction from 31.59 inches at Chatham to 39.01 inches at Stratford.

Sites Used in the Study

The daily streamflow generation is to be applied simultaneously to three sites (Figure 1):

1. Adelaide station (02GE002) on the North Branch Thames River
2. Ealing station (02GD001) on the South Branch Thames River
3. Byron station (02GD003) below the confluence of the North and South branches (Thames River).

The historical records available for the three stations are in the form of sequences of discrete values of average daily streamflows at each gauging station. Table 1 in Appendix I gives a summary of available data for these stations.

The eight-year period 1923-1930 was chosen as the concurrent period of record for the three stations, the historical data for which represent unregulated flow sequences. It was recognized that a sample of streamflow data of this size is a small one; however, a decision was made to attempt to generate synthetic data on the basis of the above-stated availability of data and for the additional two reasons:

i - The available concurrent records of regulated flows cannot be used because the extent of regulation has not been well-documented and therefore the effects cannot be defined.

ii - The comparisons of the means and standard deviations of annual flows for the eight-year period to those for longer, non-concurrent, unregulated periods available for stations 02GD001 and 02GD003 (Table 2) were within an acceptable range and no adjustments were considered necessary to the parameters for the shorter period.

LITERATURE REVIEW OF DAILY STREAMFLOW SIMULATION TECHNIQUES

General

In hydrologic studies, river discharge, as a variable observed or recorded in time, can be considered as an unidimensional (time), stochastic hydrologic process. Sequences of streamflow data characterized by statistical properties are called stochastic sequences (processes) and as they are observed or recorded in time, they constitute a time series(12). Hydrologists are often handicapped because of the short sequences of available flow data. As a solution to this problem, hydrologists during the past decade have devised various data simulation techniques which may be applied to historical records to obtain hypothetical sequences of data which preserve the significant statistical properties of the historical data. All these simulation techniques are mathematical models that make use of the computer and programming logic. In this sense, simulation is a tool that combines theory, data, and programming logic to express in mathematical terms the pertinent elements of a complex hydrologic problem (10).

Daily Streamflow Simulation Models

In the past, stochastic models have been applied to annual and monthly streamflow data, obtaining results compatible with historical flows (Thomas and Fiering (8), Beard (1), and others). Extensions of these models for generating daily flows were first attempted by Halter and Miller (5), who developed a daily flow simulator using a linear regression model which generated 30 flows each month, using as its basis the mean monthly flow and the standard error of the monthly flow. The simulated hydrographs, however, were not adequate because the serial correlation among previous flows was not incorporated into the generator; however, flows during recession periods were considered to be exponentially correlated.

At the International Hydrology Symposium (Fort Collins, 1967), Beard (2) presented a single-station model based on generation of monthly streamflows and subsequent allocation of the monthly total amounts to each day. The monthly streamflow generator used by Beard was that developed in the Hydrologic Engineering Centre of the Corps of Engineers (4). Beard's model thereby provided a realistic representation of daily flows which met the four following specifications:

i - the generated flows must maintain the expected mean of the observed flows,

ii - the variance about the monthly mean must be comparable to that of the historical record,

iii- discontinuity between daily flows at the end of one month and the beginning of the next month must be avoided,

iv - the generated flows must be properly related in time(7).

Payne, Neuman and Kerri (6), subsequently introduced a daily streamflow generation suitable for an arid region, in which they suggested the re-arrangement of the times of occurrences of flood peaks within an annual cycle, in order to preserve the pertinent properties of the basin hydrograph. In all other respects, the Payne et al model was the same as Beard's model. The operational device as suggested by Payne et al, however, appears unsuitable for humid areas where there are usually two to five peaks per month during the rainy season. Under such varying conditions, the re-arrangement of the historical record would likely pose a serious problem. The re-arrangement of the flow peaks to occur at one or two specific dates within one month, or a season, destroys the time sequence, and thereby reduces the accuracy of the response of an existing system to potential non-historical flow sequences, because the intervals between consecutive peaks or lows has been artificially fixed.

In 1972, Environment Canada (7), introduced a multi-site daily streamflow generator for the Fraser River basin. This generator used basically the same technique as described by Beard and incorporated the following two additional specifications for a multi-site model:

i - synthetic daily flows at different sites are properly related to one another in terms of location,

ii - the hydrologic lag-time between upstream and downstream stations is preserved.

The Fraser River model also uses synthetic monthly streamflow as an input. These are generated using the technique as developed by Young and Pisano (11).

It should be noted that,based on observed, historical records, statistical relationships can be derived for daily flows to allow the re-generation of the observed record; however, in order to generate sequences of daily streamflows for periods greater than the historical record, synthetic monthly means covering the same period for which daily flows are to be generated are required, as input to the daily generation program.

In reviewing the various techniques of simulation, it was concluded that Beard's model likely represents the current state-of-the-art. Beard's original model, however, was designed for a single-station problem. Therefore, Beard's technique, as modified by Environment Canada for a multi-site study in the Fraser River basin, extended to suit the Thames River, was considered to be most effective in achieving the objectives of the proposed study.

BEARD'S DAILY STREAMFLOW MODEL

The daily streamflow generation as used by Beard consists of a second order Markov chain, using standardized variates. For Markovian models, such as Beard's, each value of the process at time t depends on its immediate, previous values at times t - 1, t - 2......, and on an independent random component. In mathematical notation this can be written as follows:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \ldots\ldots + a_n X_{t-n} + E_t \quad \ldots(1)$$

where

X_t = the present value of the process,

$X_{t-1} \ldots X_{t-n}$ = the past values of the process,

$a_1 \ldots a_n$ = coefficients,

E_t = a random component with a certain probability distribution or pattern.

The parameter n gives the order of the process (i.e. order 2 in Beard's model where the present flow value is affected by flow values for the two previous days) and also represents the measure of its memory. Memory applies to the span of time during which an event has meaningfully affected the events following.

The generation of daily flows in Beard's model is accomplished by two computer programs, an analysis program and a generation program.

In the analysis program, daily flows are first classified by calendar month, and their frequency characteristics (logarithmic mean, standard deviation, skew coefficient, serial correlation coefficient, and the correlation coefficient between flow logarithms of each day and the second antecedent day) for each calendar month are determined from observed data.

Standardized variates are obtained by subtracting the mean logarithm from the daily flow logarithm, dividing this value by the standard deviation, and transforming to

normal by use of the following equation, which approximates the Pearson Type III distribution:

$$t^{*} = \frac{6}{g}\left(\left(\frac{g}{2}K + 1\right)^{1/3} - 1\right) + \frac{g}{6} \qquad \ldots(2)$$

where

t^{*} = a normal standard deviate,
K = a Pearson Type III standard deviate,
g = the skew coefficient.

The analysis program proceeds with computing the linear regression coefficients of the standard deviation of daily flow logarithms within each month of the record to logarithms of total flow for each month.

The generation program of Beard's model generates for each day standardized variates conforming to the serial correlation observed in the recorded data for that calendar month, using the following equation:

$$X_{i+2} = b_1 X_{i+1} + b_2 X_i + \sqrt{(1-R^{-2})} \cdot X_{\gamma} \qquad \ldots(3)$$

where

X_i, X_{i+1}, X_{i+2} = standardized variates for successive days,
X_{γ} = a random standardized variate,
b_1, b_2 = regression coefficients,
R^{-2} = the determination coefficient for the regression equation.

The generation program obtains the logarithm of the monthly mean flow as a first estimate of the mean logarithm of daily flows to be generated. It is noted that the generation program requires a set of monthly means (logarithmically transformed) to initiate the generation of daily sequences. If the available monthly means of the historical record are used as input to the generation program, along with the frequency characteristics of the same record obtained from the analysis program, then the generation program will reproduce (theoretically) the historical daily sequence. This characteristic of the program is used to check the ability of the model to reproduce the historic hydrograph of a particular year or years of the record for visual comparison.

In order to generate sequences of daily streamflows that are longer than the available observed record, but are statistically similar to the historical record, a set of synthetic monthly means of the same length is needed as an input to the generation program. As was stated earlier, Beard's model utilizes the monthly streamflow generator developed in the Hydrologic Engineering Centre of the Corps of Engineers, to obtain such synthetic monthly flow sequences.

The daily generation program transforms the standardized variates to logarithms of daily flows by use of the following approximate transform equation for the Pearson Type III distribution:

$$K = \frac{2}{g}\left(\frac{g}{6}\left(t^* - \frac{g}{6}\right) + 1\right)^3 - \frac{2}{g} \qquad \ldots(4)$$

which is simply another form of Equation (2). The logarithms are converted to flows which are added to obtain a monthly total. The sum includes random components and is therefore somewhat different from the desired total entered from the monthly streamflow model. A ratio of these totals could be applied to each daily streamflow, but this would ordinarily result in discontinuities at the end of each month. In order to smooth these discontinuities, Beard's model preserves the random numbers used for that month and generates new daily values on the basis of a temporary monthly total. The latter is obtained by multiplying the desired monthly total by the ratio of the desired total to the generated total, for the first pass.

The flows resulting from the second pass generation are very nearly the total of the desired monthly amount. These flows are therefore multiplied by a ratio to obtain the exact monthly total, so that a serious discontinuity at the end of each month is not created.

GENERATION OF MONTHLY STREAMFLOWS FOR THE THAMES STATIONS BASED ON THE YOUNG AND PISANO MODEL

The daily streamflow generator, as originally used for the Fraser River study, is dependent on the existence of an acceptable monthly mean generation model to provide the proper input. The mathematical technique, used for the generating of the necessary monthly streamflows for the Thames stations, is based on the Young and Pisano method (11). The technique appears more suitable than others in dealing with short periods of record for generating synthetic sequences, as all months are combined for the correlation analysis, rather than treated independently (7).

The method presumes an underlying Markovian structure and considers cross-linkages or interrelationships between stations along both the time and space axes. The advantages of the method are its straightforward approach to the problem of flow generation, its computational ease and wide applicability.

The objective, in general terms, is to extract from the historical records, statistical parameter values to construct a model of the monthly historical time sequence, in the following manner:

Consider Y years of data at n sites. The data constitute an observation matrix X, of monthly values, having dimension n . t (t = 12Y). The matrix X is used to construct the model which generates S years of synthetic data.

Following the Young and Pisano principles, the computer program for this model was written by the Engineering Division, Inland Waters Directorate, Department of the Environment. The program proceeds along the following steps:

1. Computes the residuals, using the following equation:

$$r_{ij}^{K} = \frac{X_{ij}^{K} - m_{Kj}}{S_{Kj}} \qquad \ldots (5)$$

where

r_{ij}^{K} - are the residuals,

X_{ij}^{K} - are individual elements of matrix X,

S_{Kj} - are the corresponding standard deviations,

m_{Kj} - is the average monthly value (an average of the monthly averages),

K - denotes the site ($K = 1, 2, \ldots, n$),

i - denotes the year ($i = 1, 2, \ldots, Y$),

j - denotes the month ($j = 1, 2, \ldots, 12$).

Each element in matrix X is replaced by its corresponding residual to yield a residual matrix R_1 of dimension $n \cdot t$ ($t = 12Y$).

2. Transforms the historical record in matrix X, by taking $\log_{10} X_{ij}^{K}$, $\sqrt{X_{ij}^{K}}$ of each element and transforms further the new matrices into R_2, R_3 matrices.

3. Compares the matrices R_1, R_2, R_3 and adopts the R matrix with the smallest average skewness. As normal data have an expected skewness of zero, the strategy is to pick up the option which yields a minimum average skewness. Average skewness is the sum of skew coefficients at each site (row) divided by n. This initial operation was referred to by Young and Pisano as the finding of the MST (Minimum Skewness Transformation).

4. Computes the means and standard deviations of the normalized data, matrix R, for each calendar month of each site and then standardizes the data. At this stage the data are assumed to be normally distributed with zero mean and unit standard deviation.

5. Computes the matrices Mo and M_{-1} which are the sample estimates of the variance-covariance matrix and the lag-one covariance matrix of the residuals, using the following equations:

$$Mo = RR^T \left(\frac{1}{n}\right) \qquad \ldots(6)$$

$$M_{-1} = R\bar{R}^T \left(\frac{1}{n}\right) \qquad \ldots(7)$$

where

Mo - is the variance-covariance matrix (the cross-correlation matrix),

M_{-1} - is the lag-one covariance matrix (the lag-one correlation matrix),

R - is the residual matrix,

$\bar{R}$ - is the same as R except that the rows have been shifted one time frame and the first column of R becomes the last column of $\bar{R}$,

R^T, $\bar{R}^T$ are transposed R and $\bar{R}$ matrices.

6. Sets up the generation model which is based on the following equation:

$$R_\ell = AR_{\ell-1} + Be_\ell \qquad \text{...(8)}$$

where

R_ℓ - is the ℓ^{th} column of residuals,

A - is an n x n matrix,

B - is an n x n matrix,

e - is a n x 1 vector of independent (0,1) normal deviates,

ℓ - is the time frame.

The matrix A is obtained using the following equation:

$$A = M_{-1}\, M_o^{-1} \qquad \text{...(9)}$$

where

M_{-1} - is the lag-one covariance matrix,

M_o^{-1} - is the inverse of M_o matrix.

The matrix B is obtained using the following equation:

$$BB^T = M_o - M_{-1} \cdot M_o^{-1} \cdot M^T_{-1} \qquad \text{...(10)}$$

where

B^T - is the transposed B matrix,

M^T_{-1} - is the transposed M_{-1} matrix, and all other terms are as previously defined.

The remaining analysis pertains to the solution of Equation (10). The matrix B is obtained by assuming B to be a lower triangular matrix. This assumption allows a solution for the coefficients of B.

7. Generates synthetic residuals: (1) set $R_o = 0$ (all elements); (2) calculates R_1 and succeeding values, R_ℓ, from Equation (8).

8. Destandardizes the data using the following equation:

$$X_{ij}^{K} = S_{Kj} \cdot r_{ij}^{K} + m_{Kj} \qquad \ldots(11)$$

where all terms are as defined for Equation (5).

9. Uses the inverse of the adopted normality transform to produce sequences of synthetic monthly data.

GENERATION OF DAILY STREAMFLOWS FOR THE THAMES RIVER STATIONS BASED ON THE FRASER RIVER MODEL

The mathematical techniques employed in the generation of daily streamflows for the Thames River stations are basically the same as those employed in the Fraser River basin study by the Engineering Division, Environment Canada.

The generation of daily streamflows is accomplished by two computer programs. The first program is restricted to an analysis of the historical daily streamflows and the computation of statistical parameters of the historical data, as are required in the generation procedure. The second computer program employs these parameters to generate synthetic daily streamflow data for a specified number of years for which synthetic (from Young and Pisano monthly generation) or historic monthly flows are available. The synthetic daily flows were generated for the calendar months of May to November (inclusive), at three sites simultaneously (the Adelaide, Ealing and Byron stations).

The Historical Data Analysis Program

The program reads the station identification and location information for the three stations in question: (Byron-02GE002, Ealing-02GD001 and Adelaide-02GD003); it also reads the eight years of concurrent historical data (1923-1930) for each. The data analysis is performed for one month at a time and proceeds as follows:

1. Computes the monthly means ($\overline{X}$), standard deviations (S) and the monthly totals for the monthly flows,

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad (i = 1, 2, 3, \ldots n) \qquad \ldots(12)$$

$$S = \frac{1}{n-1} \sum_{i=1}^{n} X_i^2 - \frac{n}{n-1} \overline{X}^2, \quad (i = 1, 2 \ldots n) \qquad \ldots(13)$$

2. Transforms daily flows to $\log_{10}$ values.

3. Computes the monthly means, totals and standard deviations of transformed data.

4. Regresses the standard deviations of the transformed data on their monthly totals for each station and stores the regression coefficients and constants.

5. Computes long-term means, standard deviations and skew coefficients ($\hat{\gamma}$) by month for the transformed data,

$$\hat{\gamma} = \frac{\sum_{i=1}^{n} X_i^{\,3} - 3\overline{X} \sum_{i=1}^{n} X_i^{\,2} + 2n\overline{X}^3}{n\left(\frac{1}{n} \sum_{i=1}^{n} X_i^{\,2} - \overline{X}^2\right)^{1.5}}$$

$$(i = 1, 2, 3 \ldots n) \qquad \ldots (14)$$

6. Standardizes the transformed data by subtracting the mean logarithms from the daily flow logarithms and dividing the result by the standard deviation. (The standardized variates were considered to be log-normally distributed and the normalization step used in Beard's model was omitted from the Thames River application).

7. Performs a regression analysis for each gauging station. Daily data are regressed on flow in two antecedent days and on the lagged flow at the downstream station, 02GE002.

8. Punches out statistical parameters of the historical data to be transmitted to the data generation program.

Input Data

Card input data to the analysis program are in the following order:

1. Gauging station information cards.

2. Magnetic tape input consisting of daily streamflow data at the three sites: Byron, Adelaide and Ealing. Data are in the Water Survey of Canada format 67-002.

Output Data

The historical data analysis program prints out and punches the following information:

1. Standard deviations and long-term means of May mean flows in $\log_{10}$ form for the three gauging stations, (initializing data).
2. Gauging station information.

3. Regression constants for the three gauging stations derived from the regression analysis between historical monthly total flows and the standard deviations of daily data, for each month, about the monthly mean.

4. Regression coefficients for the three stations derived in the regression analysis outlined under item 3.

5. Regression coefficients for the seven months and three stations derived in the regression analysis between daily flows at the dependent (Byron) and independent stations (Ealing and Adelaide).

6. Regression constants for the seven months for the three stations derived in the regression analysis outlined under item 5.

7. The standard errors derived in the regression relationship between daily flows at the dependent and the independent stations.

8. Average standard deviations in $\log_{10}$ units of daily flows about the monthly mean for each month at each station, as derived from the historical record.

Daily Streamflow Data Generation Program

The generation program produces a specified number of years of streamflow hydrographs over the required period of May to November, at the three gauging stations in the Thames River basin. These flow sequences are based on the statistical information derived from the historical daily streamflows at each of the three stations. The generation program accepts the statistical parameters prepared during the analysis of historical data, and a set of synthesized monthly means (see generation of synthetic monthly means) and proceeds as follows:

1. Reads the average standard deviations of daily flows about the monthly means for each station in $\log_{10}$ units.

2. Reads the monthly mean values as a first estimate of the mean daily flows to be generated for all stations being considered.

3. Computes the monthly totals using the following equation:

   monthly total = mean X number of days in month.

4. Transforms the monthly totals and means to $\log_{10}$ values.

5. Defines the flow values for the two antecedent days (29th and 30th of April, in this case) to initiate annual flow sequences. The starting values for May are set as May monthly means for that year, standardized and normalized with respect to the long-term statistics for the month.

6. Defines the total lag for the downstream station used in the regression analysis.

7. Sets the monthly standard deviation at each location to be equal to the average value of the monthly standard deviations obtained in the analysis program.

8. Generates for each day of each month, flow sequences in the form of standardized variates using the following equation:

$$X_{i+2,n} = b_1 X_{i+1,n} + b_2 X_{i,n} + b_3 X_{i+2+\ell_1,n-1} + b_4 X_{i+2+\ell_2,n-2} + \sqrt{1-\bar{R}^2} \,.X_r \qquad ...(15)$$

where

$X_{i,n}$, $X_{i+1,n}$, $X_{i+2,n}$ - are standardized variates for successive days at station n,

$X_{i+2+\ell_1,n-1}$, $X_{i+2+\ell_2,n-2}$ - are standardized variates at the stations downstream of station n on day i + 2 + ℓ, where ℓ_1, represents the lag time between n and n-1, and ℓ_2 the lag time between n and n-2 (the lag time for the Thames River study was assumed to be equal to zero because of the proximity of the stations to each other),

$\bar{R}^2$ - is the determination coefficient for the regression equation,

b_1, b_2, b_3, b_4 - are regression coefficients,

X_r - is a random standardized variate.

The similarity between Equation (15) and Equation (3) in Beard's model is obvious, the only difference being that Equation (15) takes into consideration the effect of the flow at stations n-1, n-2 on the flow at station n.

9. Transforms the sequences of standardized variates into logarithms of daily flows by multiplying each standardized variate by the log value of the standard deviation and adding to that the log value of the mean. These logarithms are then converted to naturalized flows, which are added to obtain a monthly total.

   The final monthly total is computed in the program through two passes (iterations) as described in Beard's model. When the first iteration is completed, a new target total is set for the second iteration. After the second iteration is completed, the program performs a correction by multiplying each generated daily streamflow by a correction factor. The correction factor equals the ratio between the desired monthly mean and the generated monthly mean inputs.

10. Checks on month-end discontinuities. If the discontinuity is within an acceptable limit, the sequence of generated data for a station and a month is accepted and stored. Otherwise, (i.e., if discontinuity is outside the acceptable limits) the flow sequences are regenerated up to ten times and the sequence with zero or minimum discontinuity is accepted and stored.

Input Data

The input data for the generation program are the statistical parameters derived in the historical data analysis program and a set of monthly means for the required number of years to be synthesized, as outlined previously.

Output Data

The program output consists of synthetic daily streamflows which are printed out and written in card-image form, three card images per month, on magnetic tape in the following format:

	Column	Format	Remarks
Card 1	1	A1	a letter "S" signifying synthetic data
	3-8	I1A2I3	a Water Survey of Canada identification code
	9-11	I3	year identification
	12-13	I2	month identification (5 to 11 for May to September)
	14	I1	the integer "1" denoting the first data card for the month
	15-75	10I6	synthetic flows for the first ten days in the month
	79-80	I6	the total number of daily flow observations in the month
Card 2	1	A1	as on card 1
	3-8	I1A2I3	as on card 1
	9-11	I3	as on card 1
	12-13	I2	as on card 1
	14	I1	an integer "2" denoting the second data card for the month
	15-75	10I6	synthetic flows for days 11 to 20 in the month
Card 3	1	A1	as on card 1 and 2
	3-8	I1A2I3	as on card 1 and 2
	9-11	I3	as on card 1 and 2
	12-13	I2	as on card 1 and 2
	14	I1	the integer "3" denoting third data card for the month
	15-80	11I6	synthetic flows for days 21 to the end of the month

TESTING OF THE GENERATED DATA

General Remarks

In synthetic data production, the generated data must preserve the significant statistical properties of the historical data. As Fiering and Jackson (3) indicate, the most that any generating scheme can promise about any statistic (mean, standard deviation, etc.) of the historical data is that the expected value of the same statistic is the desired one. The expected value of a statistic is its average value in an infinitely long sequence. However, because the generation algorithms are only used to form finite sequences of flows, the sample statistics obtained cannot be expected to be exactly equal to the historical statistics. They only tend to be near the historical statistics, and the closeness is expected to improve with the length of the generated data. Therefore, the preservation of the significant statistical properties of the historical record within the synthetic sequence implies that the expected values of these statistics are the specified values. The actual obtained value of a statistic for a finite synthetic sequence will be within theoretically anticipated sampling errors of the expected value. This qualification should be taken into consideration during the examination of the test results that provide comparisons between the means, standard deviations and coefficients of skew for individual months of the historical and synthetic data.

Results of the Synthetic Monthly Flow Data

A sequence of 1,000 years of monthly means was generated and one set of 50 years, exhibiting the most extreme values in the low flow range, was chosen for the study. Low flow extremes were considered to be of particular importance in evaluating possible critical water quality conditions in the river.

A series of tests were undertaken to check the resemblance between the monthly historic record and that of the selected 50 years of synthetic data. These tests included the following:

1. comparison between the means, standard deviations and coefficients of skewness of the historic record and synthetic data,

2. comparison between the cross-correlation matrices and the lag-one correlation matrices,

3. comparison between the auto-correlation coefficients,

4. comparison between the cyclical characteristics of both the historic and synthetic data (spectral analysis),

5. comparison between "Hurst's K" coefficients,

6. comparison between the number of changes in quartile and the number of changes in the direction of movement for both the historic and synthetic data.

Tables 3, 4 and 5 provide comparisons between the monthly means, the standard deviations of the monthly means about the long-term mean, and the average monthly skew coefficients of the historical flows and the selected set of synthetic flows. Corresponding values for both types of data in these tables are close but not identical. This is to be expected as the generation algorithm is designed to reproduce the historical statistics (in this case mainly the mean and standard deviation) only if an infinitely long sequence of synthetic data is generated. Different finite synthetic sets will have different statistics that are all distributed around the historical statistics within theoretically anticipated sampling errors. This property allows the user to choose such synthetic sets that exhibit extreme values in evaluating, for example, the effects of floods or droughts for various water management studies.

The comparison between the average means, standard deviations and coefficients of skew, as shown in Table 6 for all months of the historic and synthetic data, show closer agreement indicating that the statistics for the finite synthetic sample tend to be near the expected values of the historic statistics.

Table 7 shows the cross-correlation matrix Mo, the lag-one correlation matrix M_{-1}, and the corresponding estimates of A and B matrices for the historical data, as indicated in equations 2, 3 and 4. The matrices Mo and M_{-1} of the synthetic flows are given in Table 8. Close agreement among corresponding entries in tables 7 and 8 is evident.

Table 9 provides a comparison between the cross-correlation matrices of the historic and synthetic data on a monthly basis. These data also indicate general agreement.

The auto-correlation test (Table 10) compares the correlograms of the historical and synthetic flows. As was stated earlier, the monthly generator presumes an underlying Markovian structure of the first order. This assumption implies that the flow during any month is considered by the model to be meaningfully affected by the flow in the previous month only. As it can be seen from Table 10, the correlograms of the historical record of lags higher than one have much smaller values than those for lag one. This shows that the presumed underlying Markovian structure of order one in the monthly generator is adequate. Table 10 indicates that the lag-one correlograms of the historical and synthetic flows are close. It also shows that the correlograms of higher order compare favourably.

The cyclical characteristics of the historical and synthetic flows obtained by the spectral analysis test (Table 11) indicate general agreement.

Streamflow data are characterized by persistence. Hurst's coefficient (K) is considered to be an indicator of the degree of persistence in streamflow data. Table 12 compares the Hurst's K values for the historical and synthetic data. As can be seen from Table 12, K values for the historical and synthetic data are close, both being within the range 0.57 to 0.63.

The tests for the number of changes in quartile and the number of changes in the direction of movement (tables 13 and 14) for the historical and synthetic data indicate fairly good resemblance; however, the synthetic data show higher values. This is to be expected since 50 years of synthetic data are compared to only eight years of historical records.

Results of the Synthetic Daily Flow Data

A 50-year sequence of synthetic monthly means, generated according to the Young and Pisano method, was used to generate a 50-year sequence of daily flows. Two types of tests were undertaken to check the ability of the daily flow generator to reproduce realistic daily streamflows:

1. daily streamflows were first generated based on available historical monthly means and the generated hydrographs were visually compared with the historical hydrographs for the period 1923-1930,

2. a 50-year sequence of synthetic daily flows, as well as the historical record for the period 1923-1930 (natural flows) and for the period 1956-1970 (regulated flows), were statistically tested and the results were compared on a monthly basis. In addition, the cumulative probability distributions of the historical and synthetic minimum daily flows for stations 02GD001 and 02GD003 were compared.

Plots of the synthetic and historic hydrographs for the three stations under study are given in figures 2, 3 and 4, for the year 1930. These plots illustrate the ability of the daily flow generator to reproduce realistic flow sequences.

Table 15 provides a comparison between the means and average standard deviations of the historic record for the periods 1923-1930 and 1956-1970 and for 50 years of synthetic data on a monthly basis (May to November). It is clear from Table 15 that the values of the above-mentioned two statistics

are very close for the historic period 1923-1930 (natural flows) and the period 1956-1970 (regulated flows). This fact indicates that the regulation did not drastically change the characteristics of flow during the months of May to November. Table 15 also indicates that the means and average standard deviations for the synthetic data compare favourably with their corresponding values for the historical records.

Table 16 provides a comparison between the skew coefficients of the above-mentioned three sets of data. The comparison indicates that the synthetic data, in general, are slightly less skewed than the historical data. Table 17, however, shows that the long-term means and their standard deviations, as well as their skew coefficients, are in close agreement for both the synthetic and historical data.

Table 18 provides a list of minimum daily discharges as recorded annually at stations 02GE002, 02GD001 and 02GD003 for the period 1916-1970. In examining this table, it is evident that the extreme minimum daily discharges at stations 02GD001 and 02GD003 can reach as low as 2.0 cfs. The regulated record at station 02GD003, for the period 1955-1970, contains flows as low as 0.5 cfs. Figures 5 and 6 illustrate the cumulative probability distributions of historical and synthetic minimum daily flows for stations 02GD001 and 02GD003. Extreme values for historical and synthetic flows are either close or identical in both figures, whereas the intermediate values are lower for the synthetic data. This indicates that this particular set of synthetic flows provides a lower envelope for probable flow sequences and consequently can give the user estimates of the expected frequency and indications of the severity of critical low flow conditions.

CONCLUSION

A set of 1000 years of monthly streamflows, as well as one 50-year set of daily streamflows have been generated for the study. The statistical tests of the generated data indicate that the applied streamflow generator satisfactorily reproduced the most relevant characteristics of the historical records.

The synthetic daily streamflows that were generated provide a lower envelope for probable flow sequences.

RECOMMENDATION

For a complete analysis of the sensitivity of various water quality control alternatives, it is advisable that additional 50-year sets of synthetic daily sequences be selected. This will allow the user to obtain all the possible flow ranges, including the medium and high ranges.

REFERENCES

1. Beard, L. R., 'Use of the Interrelated Records to Simulate Streamflow', Journal of the Hydraulics Division, ASCE, September, 1965.

2. Beard, L. R., 'Simulation of Daily Streamflow, International Hydrology Symposium, Fort Collins, September, 1967.

3. Fiering, M.B. and Jackson, B. B., 'Synthetic Streamflows', American Geophysical Union, Water Resources Monograph 1, Washington, D. C., 1970.

4. 'Generalized Computer Program HEC-4, Monthly Streamflow Simulation', U.S. Army Corps of Engineers, Hydrologic Engineering Center, February, 1971.

5. Halter, A. N. and Miller, S. F., 'River Basin Planning: A Simulation Approach', Oregon State Agricultural Experiment Station, Corvallis, Oregon, 1967.

6. Payne, K., Neuman, W. R. and Kerri, K.D., 'Daily Streamflow Simulation', Journal of the Hydraulics Division, ASCE, July, 1969.

7. 'Streamflow Synthesis Studies for the Fraser River Basin', Engineering Division, Water Planning and Operations Branch, Environment Canada, Ottawa, 1972.

8. Thomas, H. A., Jr. and Fiering, M.B., 'Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation', in Mass, A., et al, The Design of Water Resources Systems, Harvard University Press, Cambridge, 1962.

9. 'Upper Thames Valley Report', Dept. of Planning and Development, Toronto, 1952.

10. Young, R. A. and Bredehoeft, J.D., 'Digital Computer Simulation for Solving Problems of Conjunctive Ground-water and Surface Water Systems', Water Resources Research, Vol. 8, No. 3, June, 1972.

11. Young, G.K. and Pisano, W.C., 'Operational Hydrology Using Residuals', Journal of the Hydraulics Division, ASCE, July,1968.

12. Yevjevich, V., 'Stochastic Processes in Hydrology', Water Resources Publications, Fort Collins, Colorado, U.S.A., 1972.

APPENDIX I

TABLE 1. Summary of Available Data for Sites Used in the Study

Station Name	Station Number	Drainage Area in mi^2	Period of Record	Remarks
Thames R. at Byron	02GE002	1200	-continuous Oct.1922-Sept.1931 -spring flood discharges Sept.1922-Aug.1955 -continuous Aug.1955-Dec.1970	records excellent. Since 1953 discharges affected by regulation at Fanshawe Dam, since 1965 at Wildwood Dam, and since 1967 at Woodstock Dam
Thames R. at Ealing	02GD001	519	-continuous May 1915-Dec.1970	records good. Since 1967 discharges affected by regulation at Woodstock Dam
Thames R. below Fanshawe (Adelaide)	02GD003	560	-continuous June 1915-Sept.1934 Oct.1935-Sept.1944 Oct.1955-Dec.1970	records excellent. Since 1953 discharges affected by regulation at Fanshawe Dam and since 1965 at Wildwood Dam

TABLE 2. Comparison of Historical Means and Standard Deviations of Long-Term and Short-Term Periods

STATION		Long-Term HISTORICAL PERIOD	Short-Term HISTORICAL PERIOD
		(1916-1960)	(1923-1930)
02GD001	Mean	474	516
	Std.Dev.	133	88
		(1916-33,1936-44)	(1923-1930)
02GD003	Mean	535	586
	Std.Dev.	136	127

TABLE 3. Comparison of Monthly Means

Means of Historical Record

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	1333	1558	4116	2561	1162	452	386	285	325	457	1183	1181
02GD001	608	710	1623	1003	461	221	171	127	146	193	460	472
02GD003	582	730	2157	1188	513	176	170	101	106	169	574	572

Means of Synthetic Data

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	988	1323	3926	2586	1317	477	363	349	273	330	983	1154
02GD001	446	630	1556	979	490	219	164	142	131	143	397	459
02GD003	423	585	2015	1187	596	193	148	135	82	95	419	531

TABLE 4. Comparison of Standard Deviations of Monthly Means About Long-Term Means

Standard Deviations of Historical Record

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	1139	1694	1261	1494	783	272	457	199	385	534	1403	1081
02GD001	512	697	427	503	282	142	173	69	161	233	543	408
02GD003	504	872	921	735	390	118	251	105	151	251	715	541

Standard Deviations of Synthetic Data

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	723	1248	1195	2019	1315	339	329	297	296	262	753	1311
02GD001	294	617	442	563	352	140	117	83	125	101	276	494
02GD003	415	585	889	977	644	165	179	151	102	90	341	630

TABLE 5. Comparison of Skew Coefficients

Skew Coefficients of Historical Record

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	0.7950	1.5860	0.6313	0.2592	0.0723	0.9998	1.7505	0.6322	1.9758	1.7421	1.3494	0.5168
02GD001	0.8954	1.2035	-0.3580	0.3448	0.4154	1.5308	1.7797	0.4459	2.1597	2.0048	1.6457	0.5481
02GD003	0.4463	1.8033	1.2220	0.1338	0.2237	0.7532	1.8548	1.3162	2.0903	1.6544	1.1365	0.4815

Skew Coefficients of Synthetic Data

Station	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
02GE002	1.1181	1.4219	1.0512	2.0727	2.4423	1.5973	2.1035	1.6222	4.0475	2.1101	1.2827	2.0977
02GD001	1.0666	1.4376	0.9672	1.4250	1.4523	1.2115	1.7305	1.0027	4.1328	1.5460	1.3887	2.2929
02GD003	1.5069	1.6330	1.7258	1.8524	2.1349	2.1493	3.3216	1.9785	4.3208	1.9102	1.0009	2.1016

TABLE 6. Comparison of the Means, Standard Deviations and Coefficients of Skew for all Months

Historical Record

Station	Mean	Standard Deviations	Coefficients of Skew
02GE002	1249	1437	1.5571
02GD001	516	558	1.4245
02GD003	586	760	1.9655
Total Skew	1.6490		

Synthetic Data

02GE002	1172	1439	2.2721
02GD001	479	533	1.8749
02GD003	534	746	2.5572
Total Skew	2.2347		

TABLE 7. Historical Correlation and Model Coefficient Matrices

M_0 - The Cross- Correlation Matrix (Variance-Covariance Matrix)

STATION	02GE002	02GD001	02GD003
02GE002	1.000	0.967	0.970
02GD001	0.967	1.000	0.922
02GD003	0.970	0.922	1.000

M_{-1} - The Lag-One Correlation Matrix (Lag-One Covariance Matrix)

02GE002	0.391	0.373	0.352
02GD001	0.396	0.400	0.353
02GD003	0.385	0.358	0.366

A Matrix

1.112	-0.221	-0.523
0.715	0.145	-0.473
0.823	-0.267	-0.186

B Matrix

0.912		
0.877	0.237	
0.891	-0.058	0.223

TABLE 8. Simulated Correlation Matrices - 50 Years of Synthetic Data

M_0 - The Cross-Correlation Matrix

STATION	02GE002	02GD001	02GD003
02GE002	1.000	0.972	0.980
02GD001	0.972	1.000	0.923
02GD003	0.980	0.923	1.000

M_{-1} - The Lag-One Correlation Matrix

02GE002	0.419	0.385	0.384
02GD001	0.426	0.410	0.394
02GD003	0.403	0.366	0.385

TABLE 9. Comparison of Cross-Correlation Matrices on a Monthly Basis

MONTH	STATION	HISTORICAL DATA 1 *	2	3	SYNTHETIC DATA 1	2	3
January	1*	1.000	0.998	0.989	1.000	0.949	0.912
	2	0.998	1.000	0.985	0.949	1.000	0.840
	3	0.989	0.985	1.000	0.912	0.840	1.000
February	1	1.000	0.985	0.993	1.000	0.980	0.948
	2	0.985	1.000	0.957	0.980	1.000	0.913
	3	0.993	0.957	1.000	0.948	0.913	1.000
March	1	1.000	0.904	0.916	1.000	0.972	0.961
	2	0.904	1.000	0.727	0.972	1.000	0.919
	3	0.916	0.727	1.000	0.961	0.919	1.000
April	1	1.000	0.990	0.988	1.000	0.952	0.978
	2	0.990	1.000	0.964	0.959	1.000	0.919
	3	0.988	0.964	1.000	0.978	0.919	1.000
May	1	1.000	0.916	0.984	1.000	0.952	0.981
	2	0.916	1.000	0.867	0.952	1.000	0.932
	3	0.984	0.867	1.000	0.981	0.932	1.000
June	1	1.000	0.917	0.841	1.000	0.971	0.970
	2	0.917	1.000	0.584	0.971	1.000	0.908
	3	0.841	0.584	1.000	0.970	0.908	1.000
July	1	1.000	0.999	0.995	1.000	0.959	0.952
	2	0.999	1.000	0.992	0.959	1.000	0.905
	3	0.995	0.992	1.000	0.952	0.905	1.000
August	1	1.000	0.960	0.921	1.000	0.958	0.984
	2	0.960	1.000	0.819	0.958	1.000	0.922
	3	0.921	0.819	1.000	0.984	0.922	1.000
September	1	1.000	0.993	0.994	1.000	0.986	0.986
	2	0.993	1.000	0.995	0.986	1.000	0.967
	3	0.999	0.995	1.000	0.986	0.967	1.000
October	1	1.000	0.988	0.994	1.000	0.938	0.942
	2	0.988	1.000	0.978	0.938	1.000	0.852
	3	0.994	0.978	1.000	0.942	0.852	1.000
November	1	1.000	0.984	0.995	1.000	0.896	0.900
	2	0.984	1.000	0.963	0.896	1.000	0.755
	3	0.995	0.963	1.000	0.900	0.755	1.000
December	1	1.000	0.985	0.991	1.000	0.960	0.981
	2	0.985	1.000	0.953	0.960	1.000	0.961
	3	0.991	0.953	1.000	0.981	0.961	1.000

* NOTE: in the final generation of the data on the computer tapes
Station 1 = 02GE002
2 = 02GD001
3 = 02GD003

TABLE 10. Comparison of Auto-Correlation Coefficients

	HISTORICAL RECORD			SYNTHETIC DATA		
	Station			Station		
Lag	1	2	3	1	2	3
0	1.000	1.000	1.000	1.000	1.000	1.000
1	0.406	0.435	0.338	0.447	0.477	0.413
2	0.012	0.015	0.010	0.101	0.136	0.063
3	-0.113	-0.123	-0.077	-0.039	-0.040	-0.056
4	-0.107	-0.159	-0.061	-0.127	-0.154	-0.131
5	-0.230	-0.270	-0.221	-0.216	-0.277	-0.215
6	-0.302	-0.322	-0.304	-0.290	-0.333	-0.277
7	-0.241	-0.272	-0.233	-0.248	-0.290	-0.234
8	-0.144	-0.172	-0.123	-0.139	-0.170	-0.124
9	-0.061	-0.056	-0.066	-0.085	-0.083	-0.076
10	0.082	0.096	0.071	0.030	0.055	0.022

Station 1 = 02GE002
2 = 02GD001
3 = 02GD003

TABLE 11. Results of Spectral Analysis

	Historical Data Station				Synthetic Data Station		
Cycle/Yr.	02GE002	02GD001	02GD003	Cycle/Yr.	02GE002	02GD001	02GD003
0.0	81	4	10	0.0	80	10	19
0.2	49	5	14	0.2	66	8	15
0.4	64	9	17	0.4	71	10	18
0.6	46	7	12	0.6	66	9	16
0.8	206	33	55	0.8	232	35	58
1.0	329	54	87	1.0	349	54	90
1.2	134	22	34	1.2	120	18	32
1.4	53	8	12	1.4	51	7	15
1.6	66	11	14	1.6	56	8	17
1.8	89	15	20	1.8	89	12	25
2.0	105	15	28	2.0	116	15	31
2.2	72	10	19	2.2	55	6	15
2.4	67	10	17	2.4	35	4	10
2.6	55	8	14	2.6	36	5	11
2.8	89	12	27	2.8	88	11	25
3.0	89	12	29	3.0	88	11	25
3.2	46	6	15	3.2	31	4	9
3.4	42	6	13	3.4	24	3	7

TABLE 12. Test for Long-Term Persistence (Hurst's K)

Station	Historic Record	Synthetic Data
02GE002	0.574	0.611
02GD001	0.609	0.618
02GD003	0.633	0.603

TABLE 13. Comparison of the Number of Changes in Quartile per 50 Years

Station	Historic Record	Synthetic Data
02GE002	250	255
02GD001	244	246
02GD003	238	259

TABLE 14. Comparison of the Number of Changes in the Direction of Movement

Station	Historic Record	Synthetic Data
02GE002	275	303
02GD001	287	296
02GD003	287	312

Table 15. Comparison of Average Means and Standard Deviations of Daily Streamflows

		Historical Period 1923-1930		Historical Period 1956-1970		Synthetic Data 50 years	
Station	Month	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
02GE002	May	1162	1036	1162	886	1322	896
02GD001	"	461	329	481	294	489	352
02GD003	"	514	544	548	497	596	475
02GE002	June	452	302	499	370	478	338
02GD001	"	222	147	224	134	218	140
02GD003	"	176	186	199	195	193	165
02GE002	July	387	382	331	172	367	326
02GD001	"	172	169	170	80	164	117
02GD003	"	170	304	108	81	147	178
02GE002	Aug.	285	205	366	231	349	217
02GD001	"	127	75	195	143	142	83
02GD003	"	101	131	119	92	135	121
02GE002	Sept.	326	242	346	196	274	246
02GD001	"	146	114	173	99	131	125
02GD003	"	106	118	117	80	82	102
02GE002	Oct.	457	213	479	394	430	262
02GD001	"	193	90	183	83	182	101
02GD003	"	169	154	216	253	165	90
02GE002	Nov.	1183	844	1019	808	1000	762
02GD001	"	460	325	325	202	397	376
02GD003	"	574	593	516	537	519	541

TABLE 16. Comparisons of the Coefficients of Skew of Daily Streamflows

Station	Month	Historical Period 1923-1930 Skew	Historical Period 1956-1970 Skew	Synthetic Data 50 years Skew
02GE002	May	4.1570	6.6283	4.8186
02GD001		2.7846	5.7123	2.7101
02GD003		4.3832	6.5280	5.7723
02GE002	June	2.2502	3.7458	2.2222
02GD001		5.1005	4.0574	2.0886
02GD003		5.1547	5.2925	2.6604
02GE002	July	9.4676	4.0811	2.6278
02GD001		8.3196	3.3525	2.3131
02GD003		9.9211	5.8258	3.8547
02GE002	Aug.	4.9345	7.5640	2.0156
02GD001		3.0688	9.3466	1.3406
02GD003		7.5945	4.9909	2.9982
02GE002	Sept.	5.9058	7.3108	4.7033
02GD001		7.1078	10.2399	4.3939
02GD003		6.5542	2.9894	5.1693
02GE002	Oct.	4.0351	7.3646	2.8319
02GD001		3.2362	3.7289	2.1656
02GD003		4.5018	8.6790	2.7428
02GE002	Nov.	3.6120	2.9645	1.9571
02GD001		2.7488	2.2322	2.1097
02GD003		4.2951	3.4186	2.2749

TABLE 17. Comparison of Long-Term Period Means, May to November, Standard Deviations of Period Means About Long-Term Mean and Skews of Same

	Historical Period 1923-1930			Historical Period 1956-1970			Synthetic Data 50 years		
Station	Mean	St.Dev.	Skew	Mean	St.Dev.	Skew	Mean	St.Dev.	Skew
02GE002	607	1157	5.9745	600	995	7.7853	589	1087	6.8440
02GD001	254	417	4.7010	254	343	7.4096	240	481	3.6716
02GD003	258	655	7.0134	260	576	7.9902	239	551	8.4569

Table 18. Annual Recorded Minimum Daily Discharges in CFS

Year	Minimum Daily Discharge Station: 02GE002	Station: 02GD001	Station: 02GD003
1916	--	38.0 Nov. 6	14.0 Sept. 6
17	--	38.0 Feb. 12	15.0 Sept. 23
18	--	2.0 Aug. 5	20.0 Aug. 7
19	--	32.0 Aug. 11	2.0 Sept. 10
20	--	48.0 Aug. 9	26.0 Sept. 27
21	--	67.0 July 7	24.0 Aug. 2
22	--	58.0 Sept. 5	12.0 Sept. 6
23	145 July 9	37.0 Aug. 6	16.0 July 31
24	109 Nov. 11	58.0 Sept. 1	19.0 Sept. 16
25	59.0 Sept. 5	42.0 Sept. 3	8.0 Sept. 2
26	95.0 July 17	40.0 July 26	18.0 July 20
27	130 Sept. 19	64.0 Sept. 26	28.0 Sept. 20
28	162 Aug. 16	80.0 Sept. 10	34.0 Sept. 10
29	106 Sept. 1	53.0 Sept. 9	27.0 Aug. 27
30	60.0 Oct. 4	50.0 July 21	6.0 Aug. 6
31	--	57.0 June 24	25.0 Sept. 13
32	--	79.0 Aug. 1	35.0 Aug. 2
33	--	55.0 Nov. 6	31.0 Sept. 7
34	--	50.0 July 2	--
35	--	31.0 Nov. 4	--
36	--	10.0 June 26	7.0 Aug. 18
37	--	36.0 Aug. 1	16.0 Aug. 2
38	--	35.0 Dec. 28	17.0 Dec. 2
39	--	45.0 Oct. 7	14.0 July 25
40	--	57.0 Aug. 12	42.0 Aug. 7
41	--	29.0 July 2	12.0 Aug. 9
42	--	46.0 Sept. 6	18.0 Aug.22
43	--	54.0 Oct. 11	32.0 Oct. 12
44	--	54.0 Aug. 16	--
45	--	49.0 Jan. 1	--
46	--	63.0 Sept. 30	--
47	--	114 Aug. 12	--
48	--	92.0 Sept. 5	--
49	--	92.0 June 27	--
50	--	107 June 11	--
51	--	114 Aug. 19	--
52	--	78.0 Oct. 27	--
53	--	58.0 Aug. 30	--
54	--	76.0 Aug. 29	5.0 Aug. 24
55	--	63.0 Oct. 2	1.0 July 29
56	199 Aug. 3	93.0 Jan. 23	38.0 Oct. 26
57	97.0 June 20	54.0 Aug. 23	38.0 Aug. 23
58	53.0 June 11	63.0 Aug. 19	4.0 Sept. 16
59	136 Aug. 4	59.0 Aug.16	19.0 Aug. 13
60	140 Sept. 13	73.0 Oct. 17	3.6 Dec. 9
61	106 Oct. 20	55.0 Jan. 24	0.60 Nov.14
62	79.0 July 14	44.5 July 16	1.6 Feb. 2
63	93.8 Aug. 26	45.6 Oct. 14	0.50 July 13
64	60.5 Oct. 30	54.6 Jan. 1	2.1 Sept.22
65	56.0 July 24	54.6 Aug.30	0.80 June 23
66	102 July 12	60.0 July 24	0.60 Jan. 16
67	285 Sept. 6	112 Sept. 20	1.4 Dec.5
68	43.3 May 30	91.0 Sept. 1	56.0 July 31
69	107 Sept. 25	88.0 Aug. 11	23.0 July 16
70	185 Sept. 13	98.5 Sept. 2	10.8 Nov. 19

Regulated Flow | Regulated | Regulated Flow

APPENDIX II

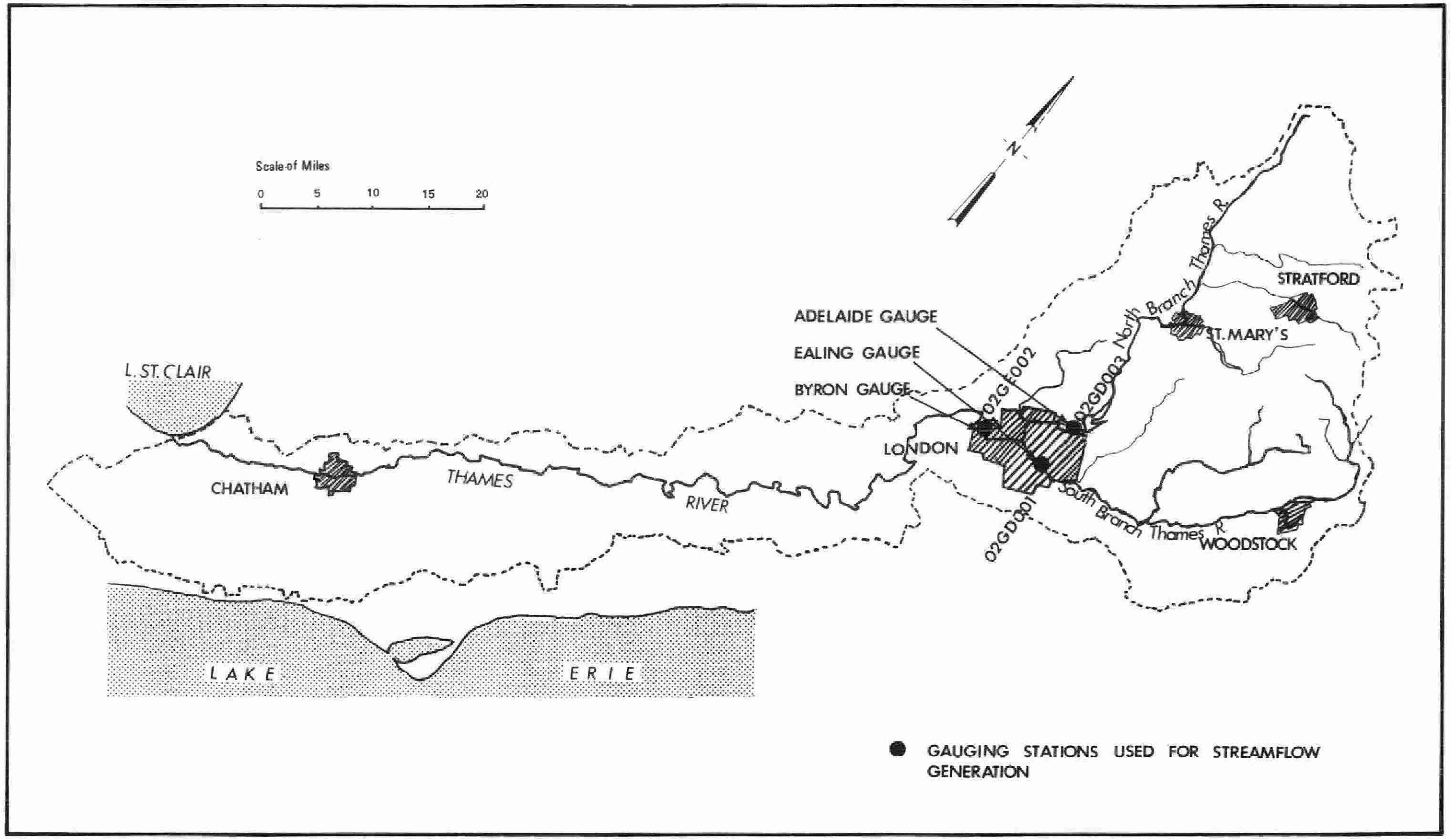

Figure I : Thames River Watershed.

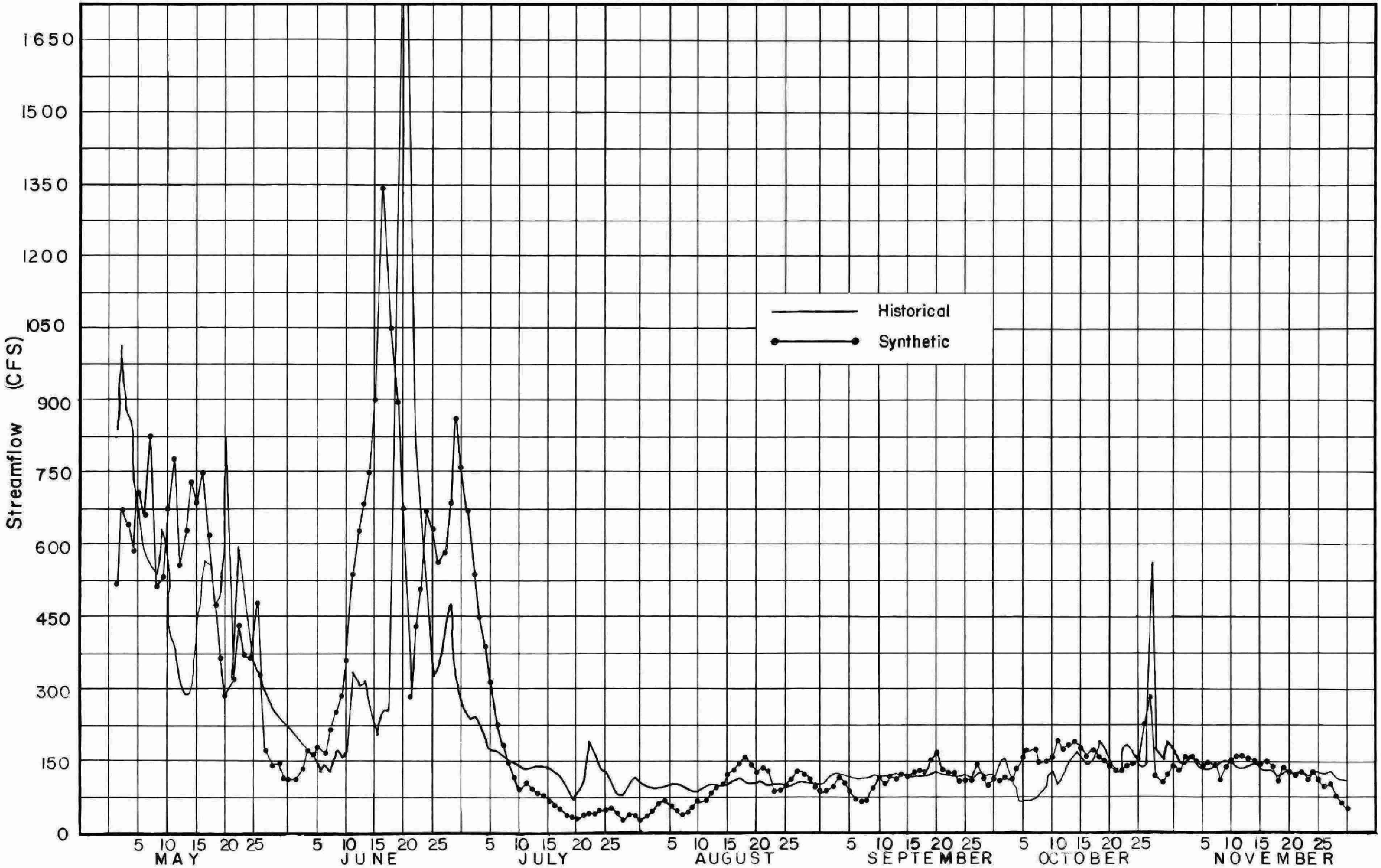

Figure 2: Hydrograph for Thames River Station: 02GE002, Year 1930.

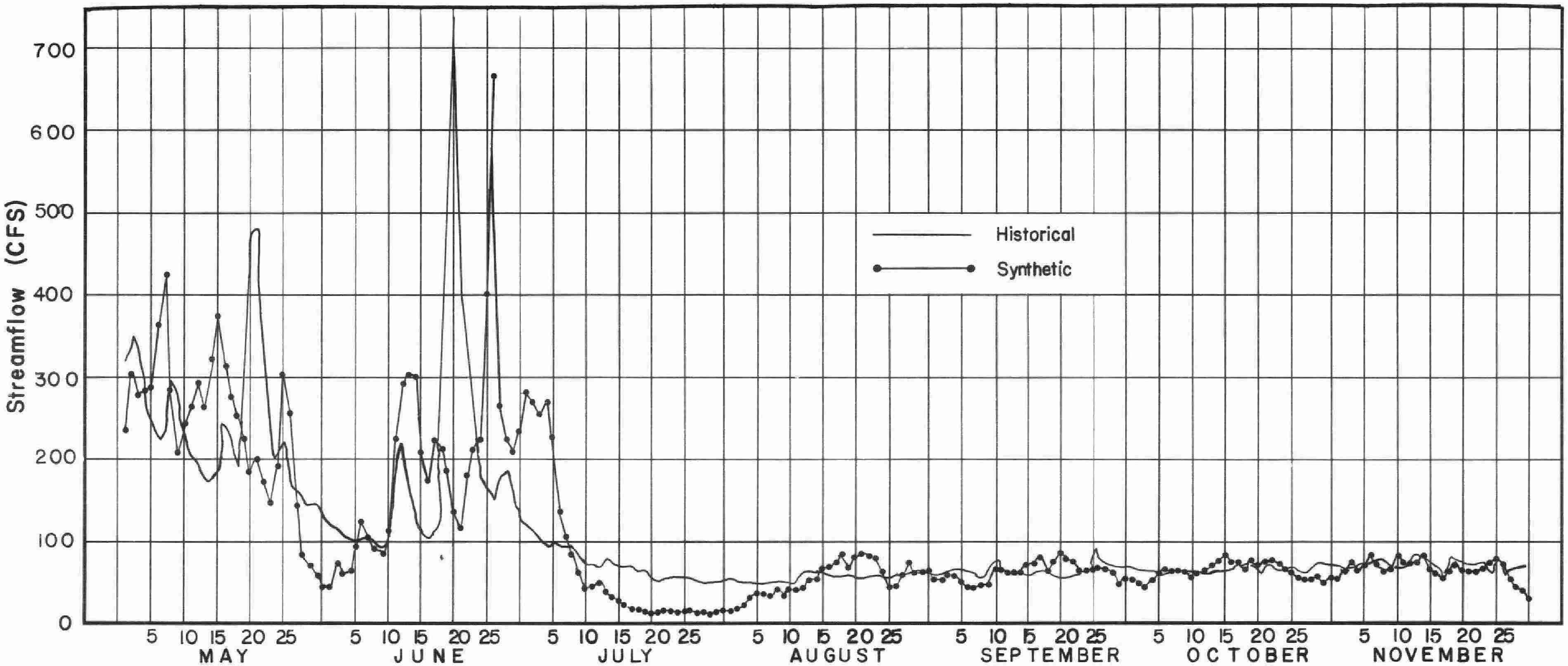

Figure 3 : Hydrograph for Thames River Station: 02GD001, Year 1930.

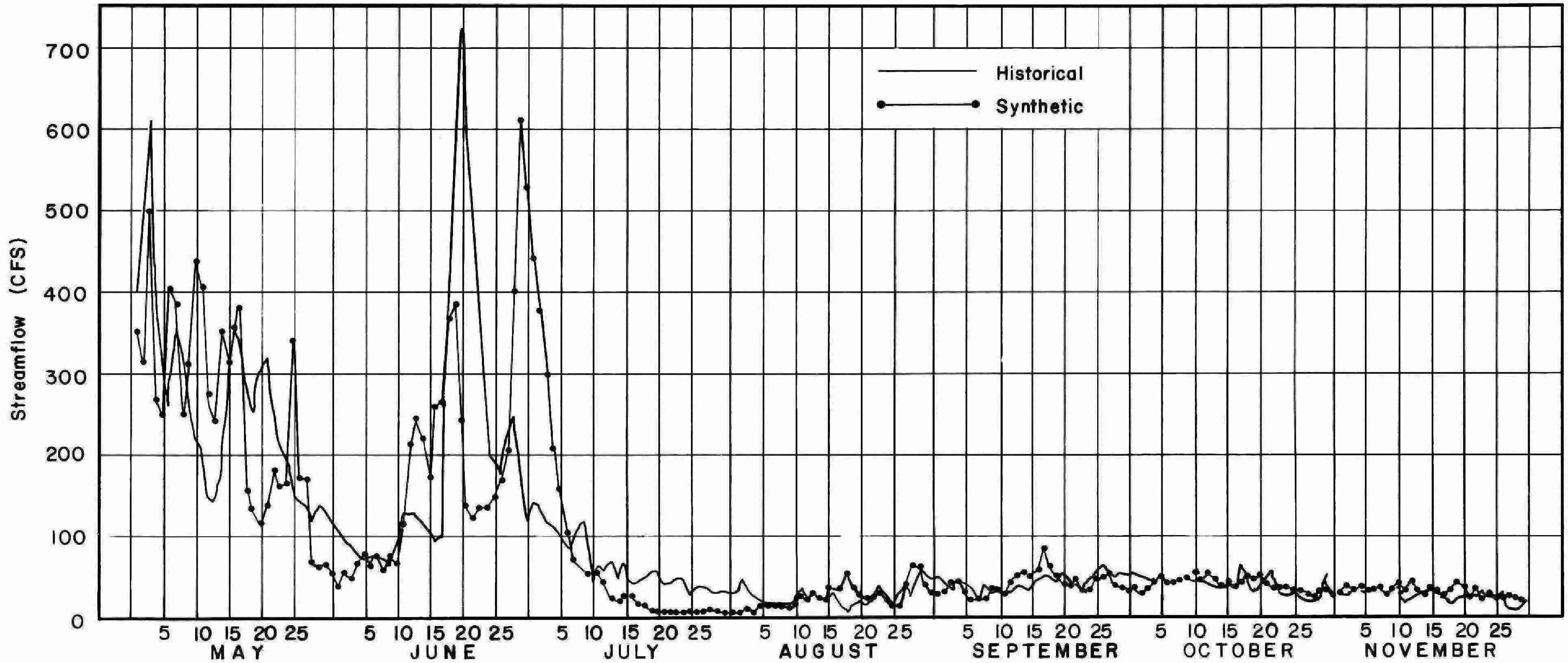

Figure 4: Hydrograph for Thames River Station: 02GD003, Year 1930.

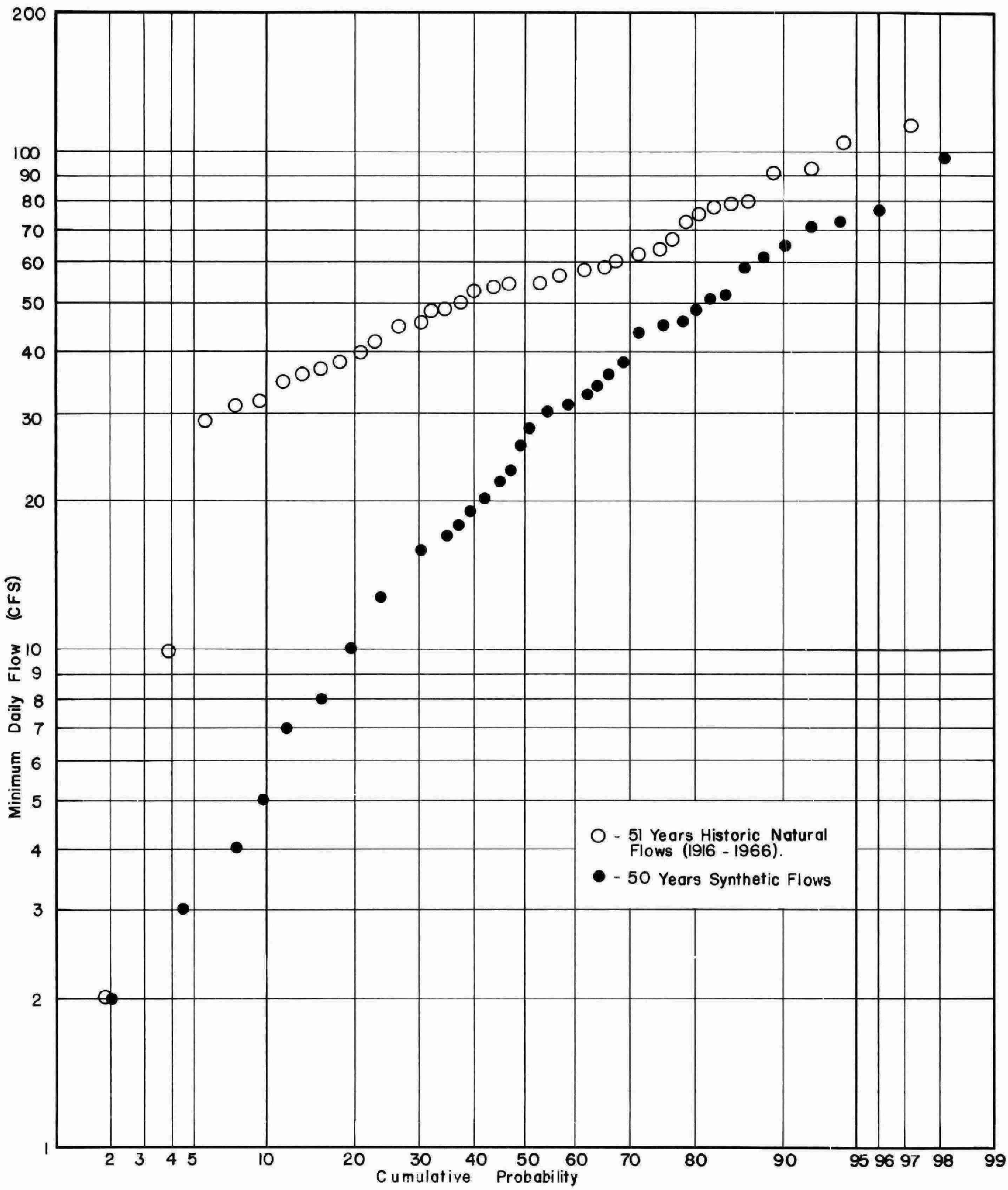

Figure 5 : Distributions of Historical and Synthetic Minimum Daily Flows, Station: 02GD001, South Branch Thames River.

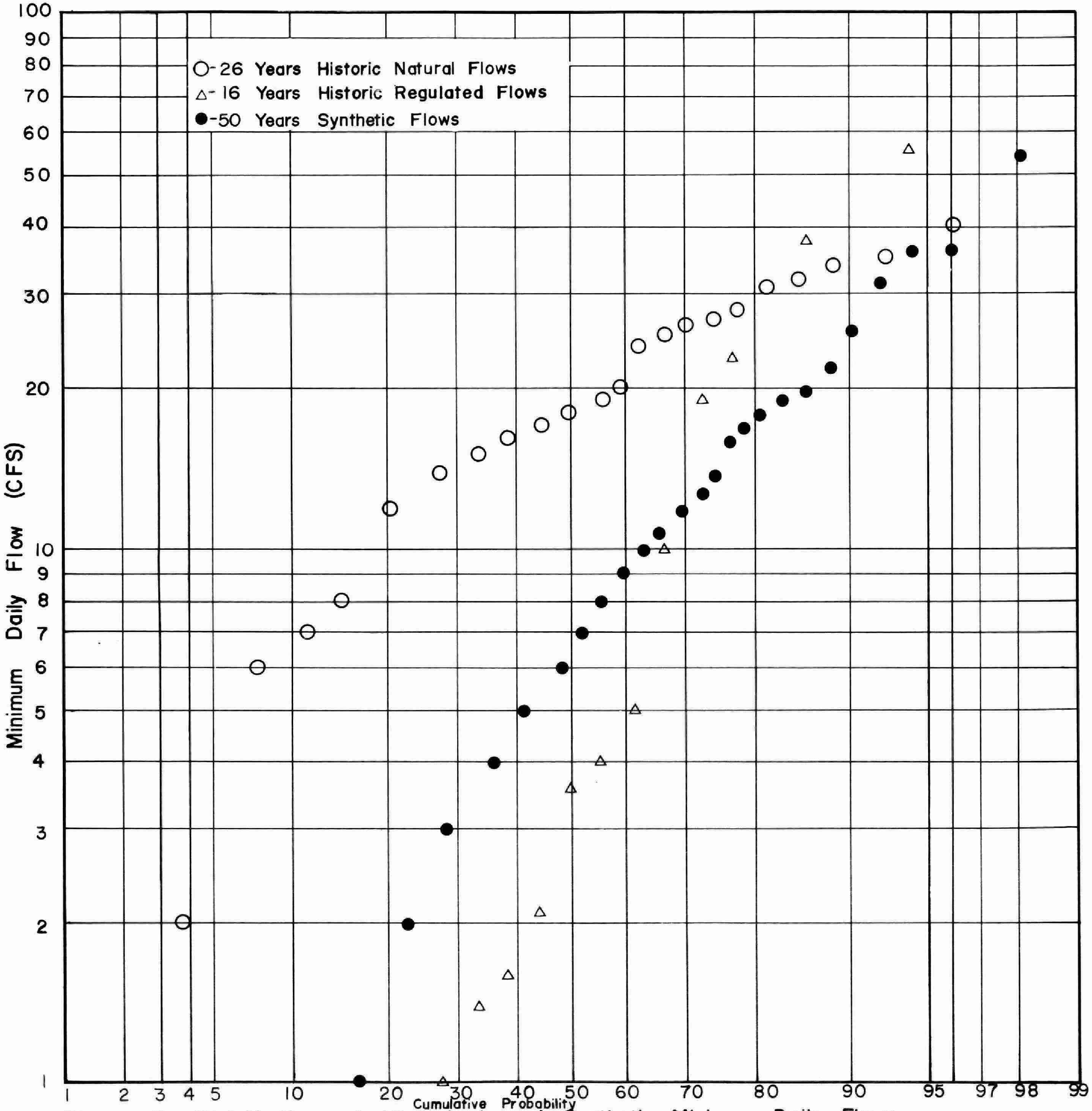

Figure 6: Distribution of Historical and Synthetic Minimum Daily Flows, Station: 02GD003, North Branch Thames River.

9693600000941O